DataFrame

Col 1 Col 2 - - - -

Row 1

Row 2

dict1 = {"A":[1,2,3], "B":[4,5,6]}

Index	A	B	C
R1	1	4	1
R2	2	5	2
R3	3	6	7

C	A	B
1		
2		
7		

Index	C	C2	States
R_1			
R_2			

	C_1	C_2	States
R_1			
R_2			

df

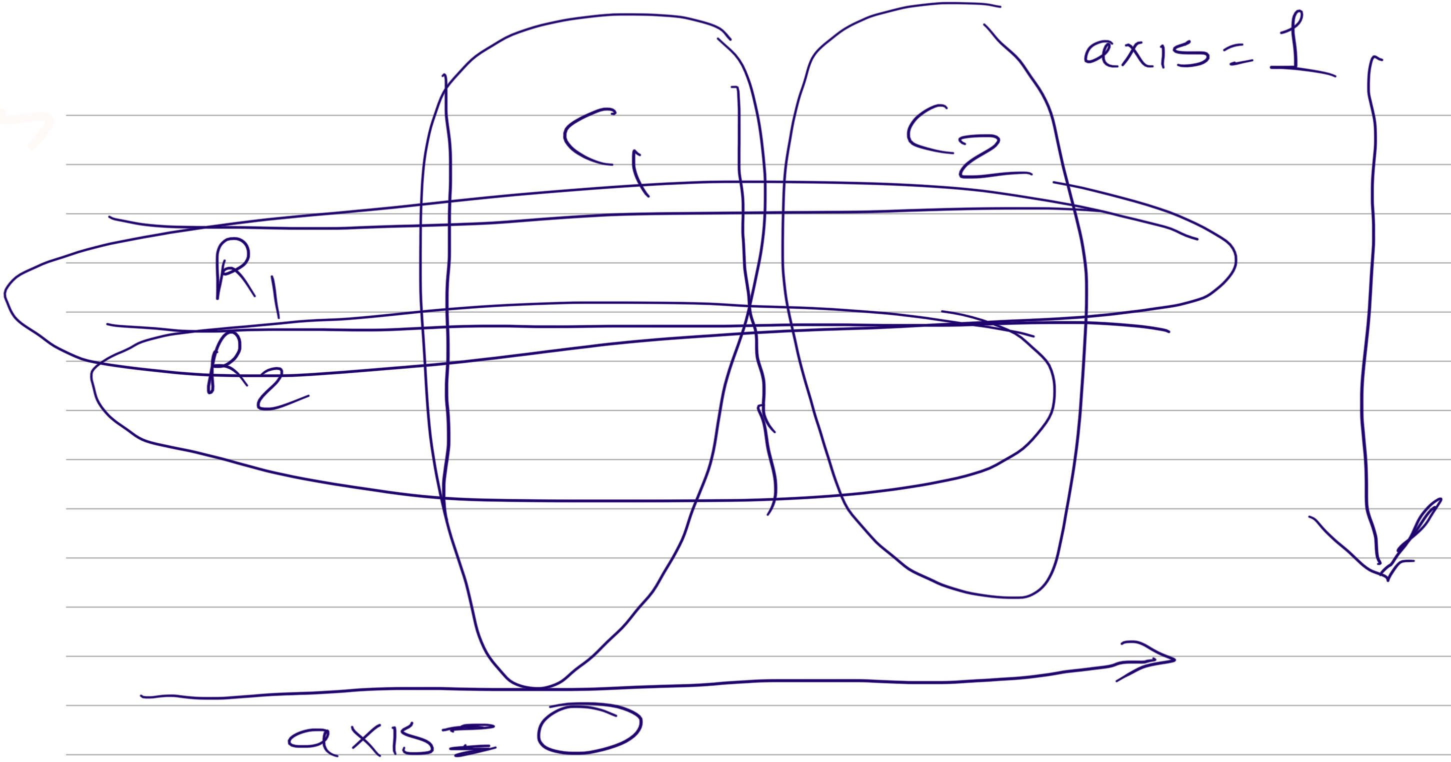
inplace = False

States

df

NaN = Not a Number

	C_1	C_2
R_1	NaN	3
R_2	2	1



df.drop('R2', axis=0)

Comp	Sales
Apple	100
Go	120
Mi	130
Go	140
Apple	150
Mi	170
Amazon	200
Amazon	

A	B	C
1	2	3
2	3	1

A	B	C
11	33	7
22	11	9

A	B	C

D	B	F

A	B	C	F

df 4 — df 5

A	B	E
A0	B0	C0
A1	B_1	C_1
A2	B_4	C_2
A3	B_{33}	C_3

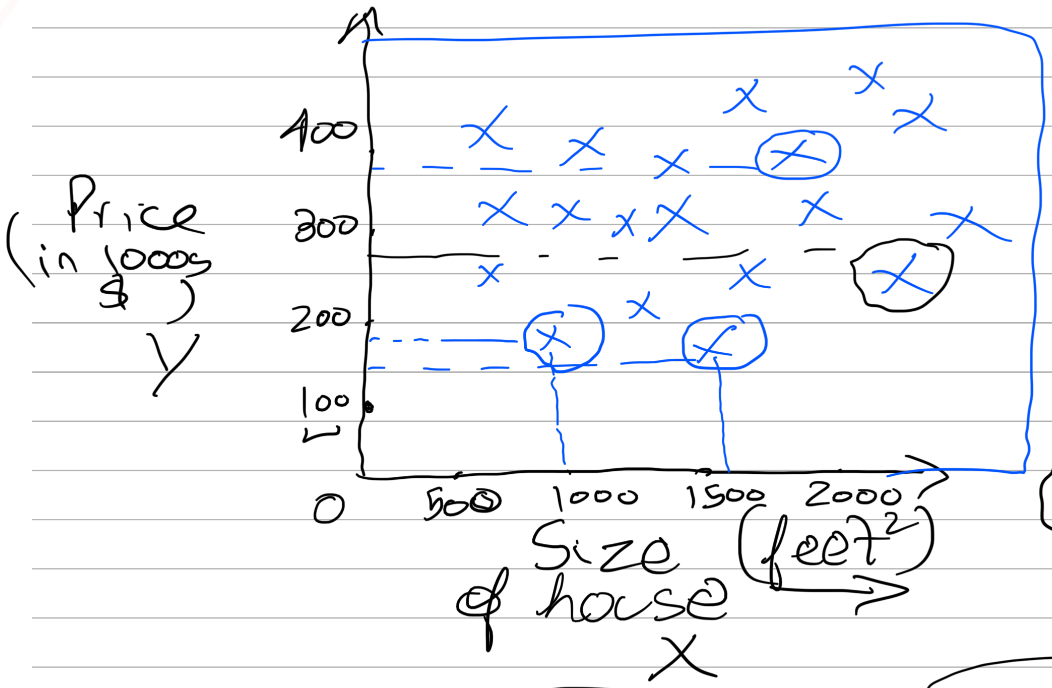
A	B	C
A0	B_0	C_0
A_5	B_4	C_{1332}
A_7	B_{22}	C_{23}
A_6	B_{33}	C_{334}

Linear

X → Y

1   2

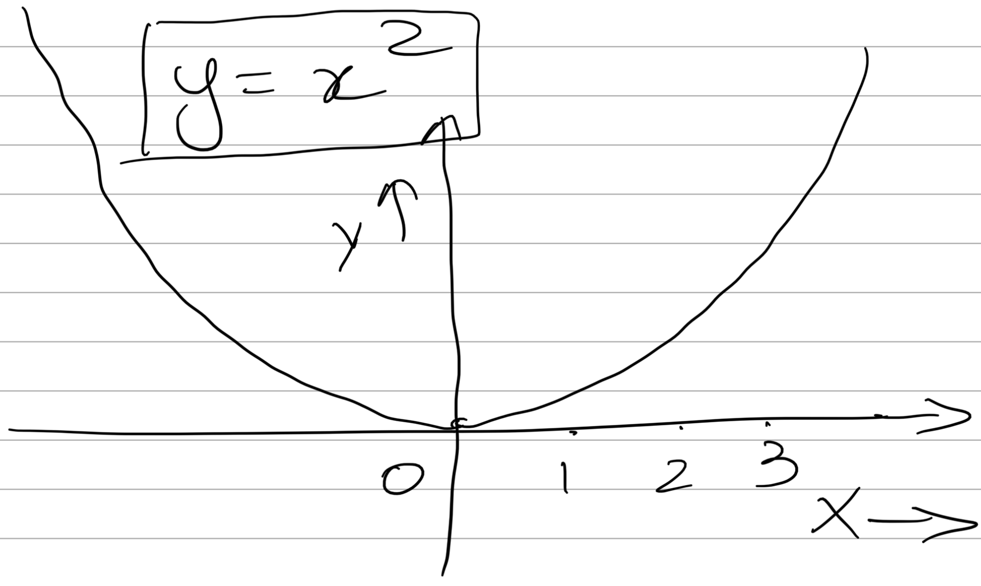
1   2.5



4000 = feet^2

1000 feet^2 = 200000

X   Y



X	Y
1	1
2	4
3	9
4	16

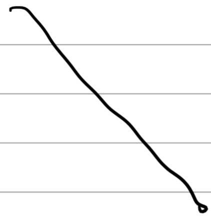
$y = x^2$

$y = 2x$

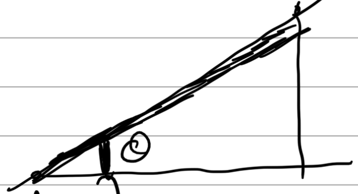
$\frac{1}{2}$ $\frac{1}{2}$

$$\boxed{y = mx + c}$$

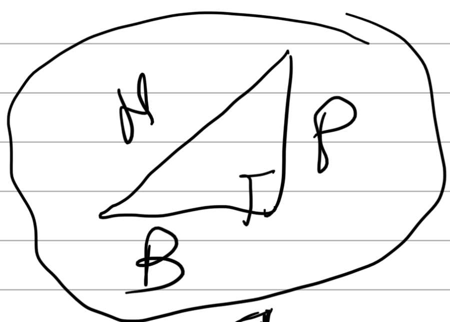
$m \rightarrow$ Slope



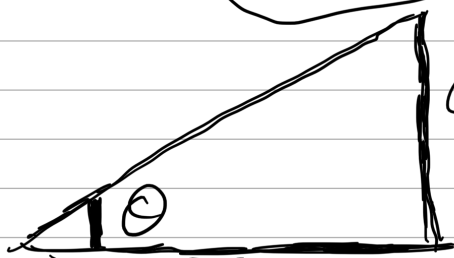
$$m = \frac{\text{Change in } y}{\text{Change in } x}$$

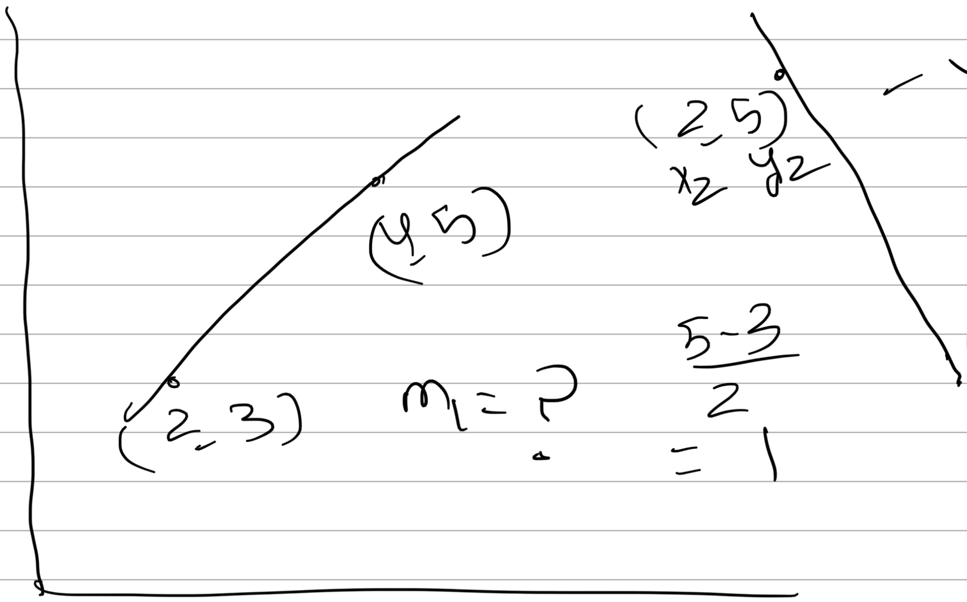


$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{5-2}{4-1} = \frac{3}{3} = 1$$



$$\tan\theta = \frac{y}{x} = \frac{P}{B}$$



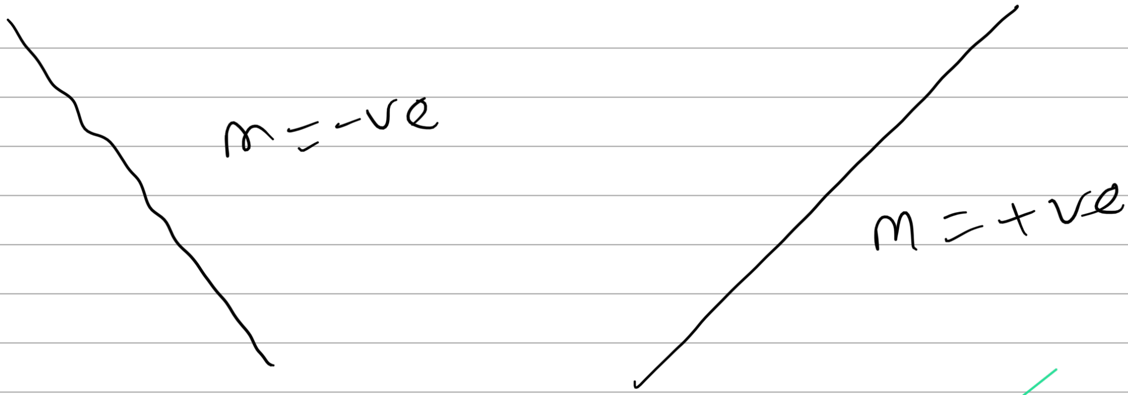


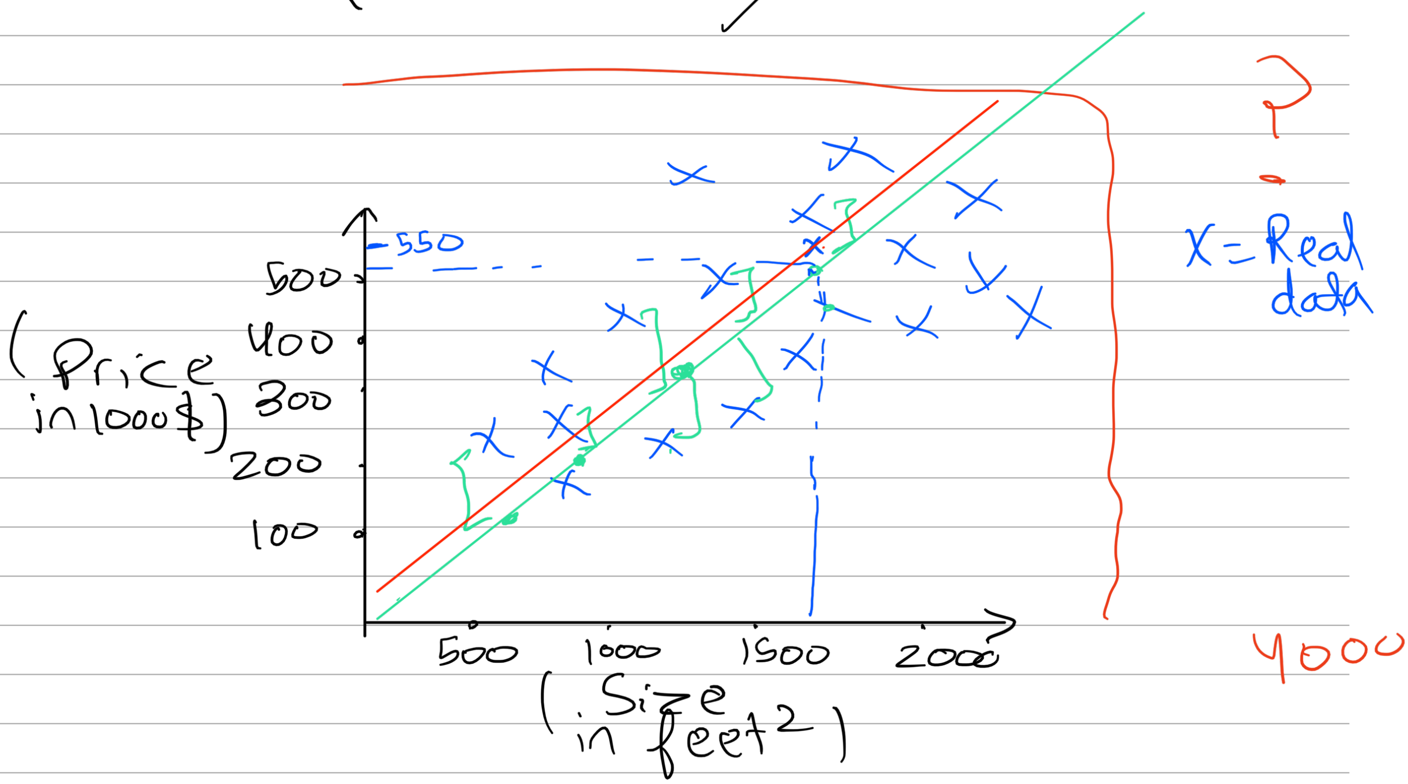
$m_1 = ?$

$$\frac{5-3}{2} = 1$$

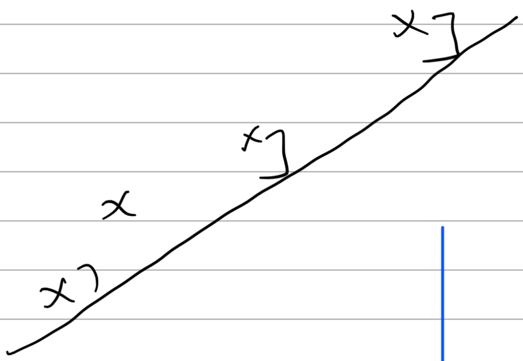
$m_2 = ?$

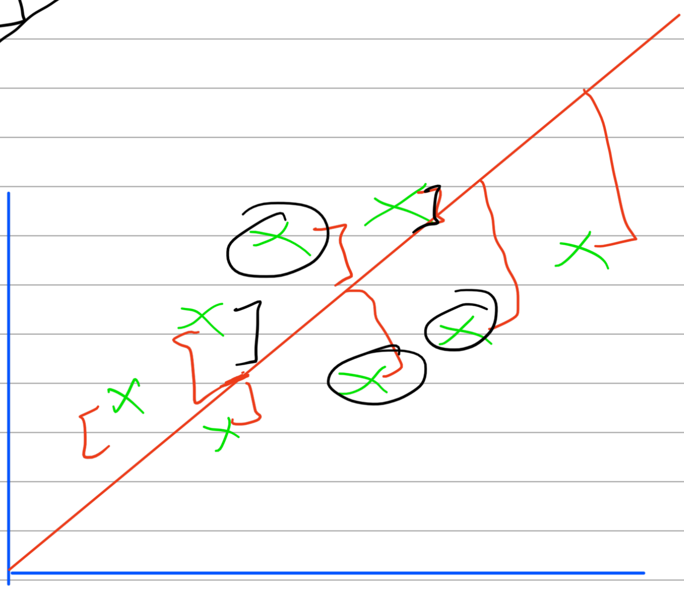
$$\frac{4}{-1} = -4$$





Error = (Real value − Predicted value)





$$\underbrace{\left(\text{Points on line} - \text{Data Points}\right)}$$

$$\left((mx_{(i)}+c) - y\right)^2$$

m = data points

$\left(m \text{ Points on the line} - m \text{ data point}\right)$

upper limit

$$\sum_{i=1}^{5} (x_i)$$

Lower

$x_i = 0$

$$\underbrace{1+2+3+4+5} = \sum_{i=1}^{5} (x_i + i)$$

$1+2+3$

Training data
↓
Linear Regression
↓
Size of House → h → Estimated Price

$$\underbrace{X \qquad\qquad y}$$

$$\text{Cost function} = \frac{1}{2m} \sum_{i=1}^{m} (\hat{y} - y)^2$$

$\hat{y}$ → Point on line

y → Real data point

m = Total data point

$$\frac{dy}{dx} \qquad \boxed{\frac{\text{Change in } Y}{\text{Change in } X}}$$

$y = x^2$

$y = x^n$

$2x^{2-1} =$

$$\boxed{y = x^n} = \frac{dy}{dx} = nx^{n-1}$$

$$\text{Cost function} = \frac{1}{2m} \sum_{i=1}^{m} (\hat{y} - y)^2$$

$\hat{y}$ → Point on line

y → Real data point

m = Total data point

$(1-2) + (2-1)$

$\boxed{-1 + 1} = 0$

$y = x^2$

$s = \frac{d}{t}$

$\text{speed} = \frac{\text{dist}}{\text{time}}$

$st = \text{dist}$

$s = 40\text{ km} \qquad t = 1\text{ hr}$

$\frac{d(\text{dist})}{dt} = \frac{d(st)}{dt} = s$

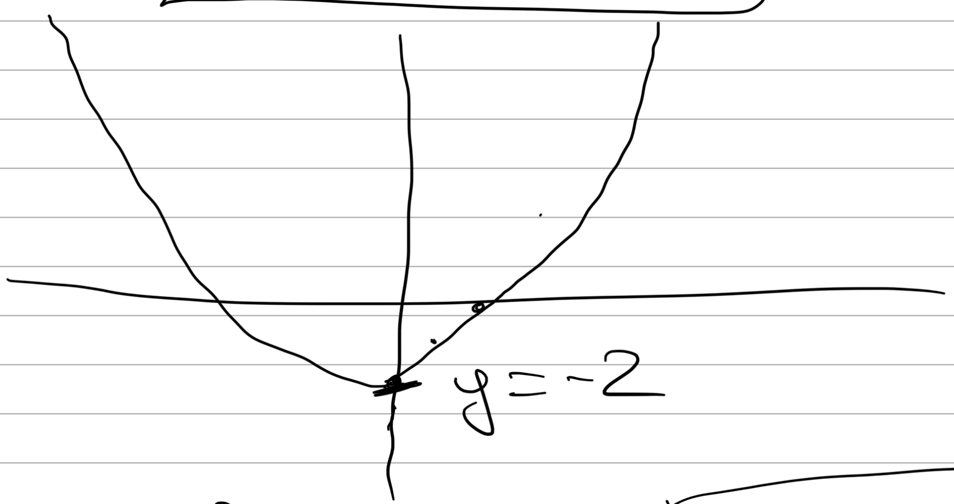
nx^{n-1}

$y = x \qquad \frac{dy}{dx} = 1$

$y = xt \qquad \frac{dy}{dx} = \frac{d(xt)}{dx}$

$= t\frac{d(x)}{dx} = t \cdot 1 = t$

$$\boxed{y = \boxed{x^2 - 2}} \qquad \frac{dy}{dx}$$



$y = x^2 - 2$

x	y
0	-2
1	-1
2	2
3	7
4	14

$y = x^2 - 2$

$\frac{dy}{dx} = 2x = 0$ $\quad x = 0$

$y = x^2 - 2$ | $x = 0$ it shows Minimum values

$y = x^2 = 2x$ $\quad \frac{1}{2}$

Cost function

$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} [h(x_i) - y_i]^2$

$h(x_i)$: Predicted Value

y_i: Real Value

θ_0: c, θ_1: m

$y = mx + c$

$y = \theta_0 + \theta_1 x$

Gradient descent

$\theta_j = \theta_j - \alpha \times \frac{d}{d\theta_j} J(\theta_0, \theta_1)$

1, 2, 3, 4, 5

$y = x^2$

$(\ (1-2) + (2-1)\ )^2$

$-1 + 1 = 0$

$y = x^2$

$x^2 =$

$= 2x$

$y = x^n =$

$\frac{dy}{dx} = nx^{n-1}$

$=$

$y = x^2$

$=$

$$y = x^2$$
$$= 2x^{2-1} = 2x$$

$$y = x^n$$
$$\frac{dy}{dx} = nx^{n-1}$$

$$(\text{Pred} - \text{Realvalue})^2$$

2

$$y = mx + c$$

$$= \frac{1}{2m} \sum_{i=1}^{m} (\hat{y} - y)^2$$

Point on line

Data point

$$\hat{y} = mx + c$$

$$c = 0$$

$$\hat{y} = mx$$

slope

$$m = 1$$

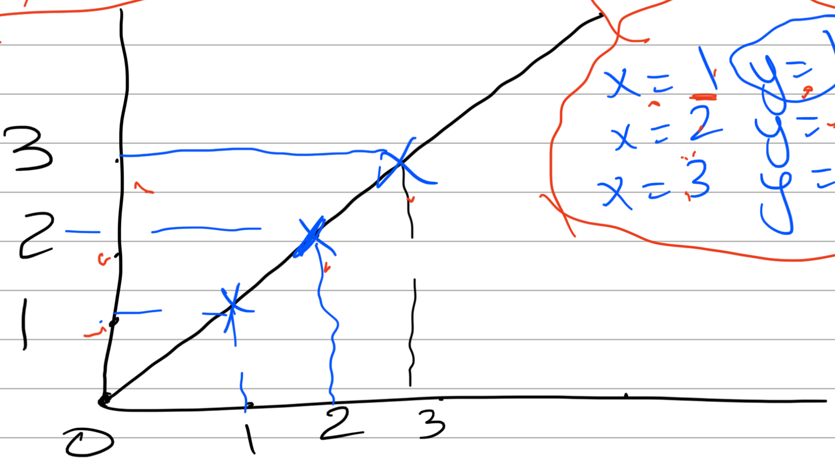
$$\hat{y} = x$$

Total data = 3

1 2 3

$y = mx$

$= \cancel{0.5x}$



$x = 1$ $y = 1$

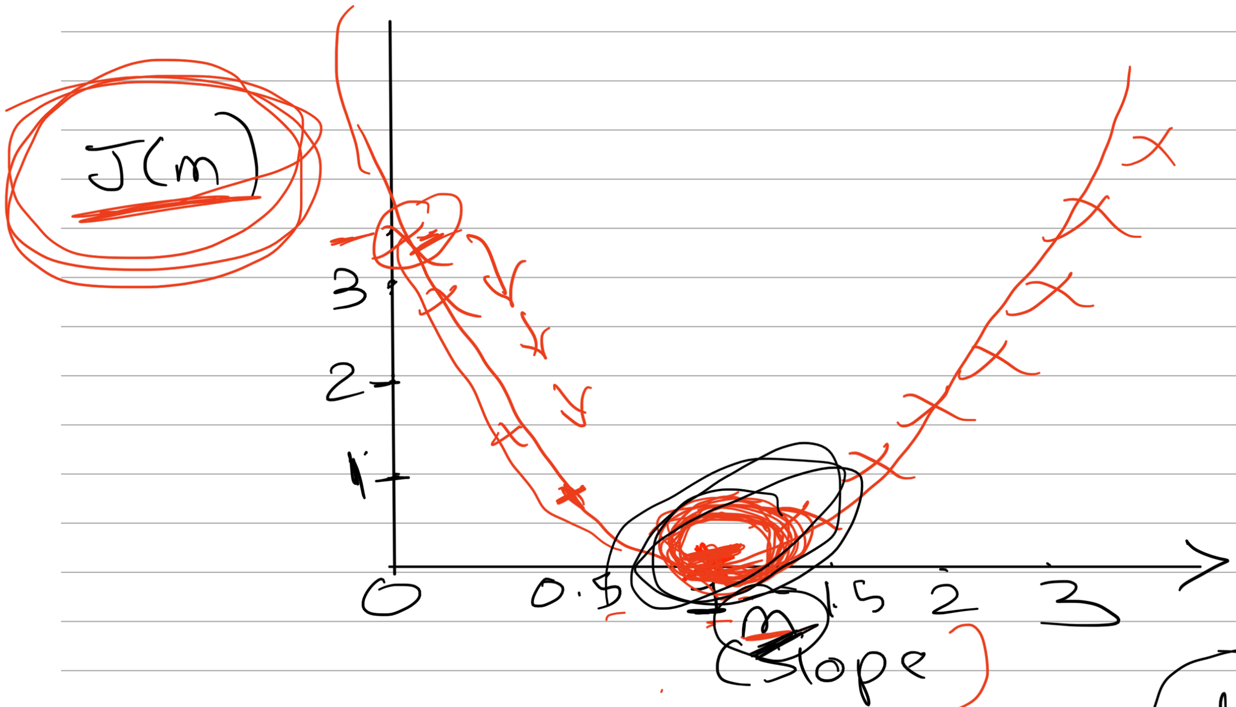
$x = 2$ $y = 2$

$x = 3$ $y = 3$

$$= \frac{1}{2 \times m} \sum_{i=1}^{m} (\hat{y} - y)^2$$

$$= \frac{1}{6} \sum_{i=1}^{3} \left((1-1)^2 + (2-2)^2 + (3-3)^2 \right)$$

$$J = \frac{1}{6} \times 0 = 0$$ Slope $\boxed{m = 1}$



$\frac{dy}{dx}$

$m = 0.5$

$\hat{y} = mx = 0.5x \quad = 0.$

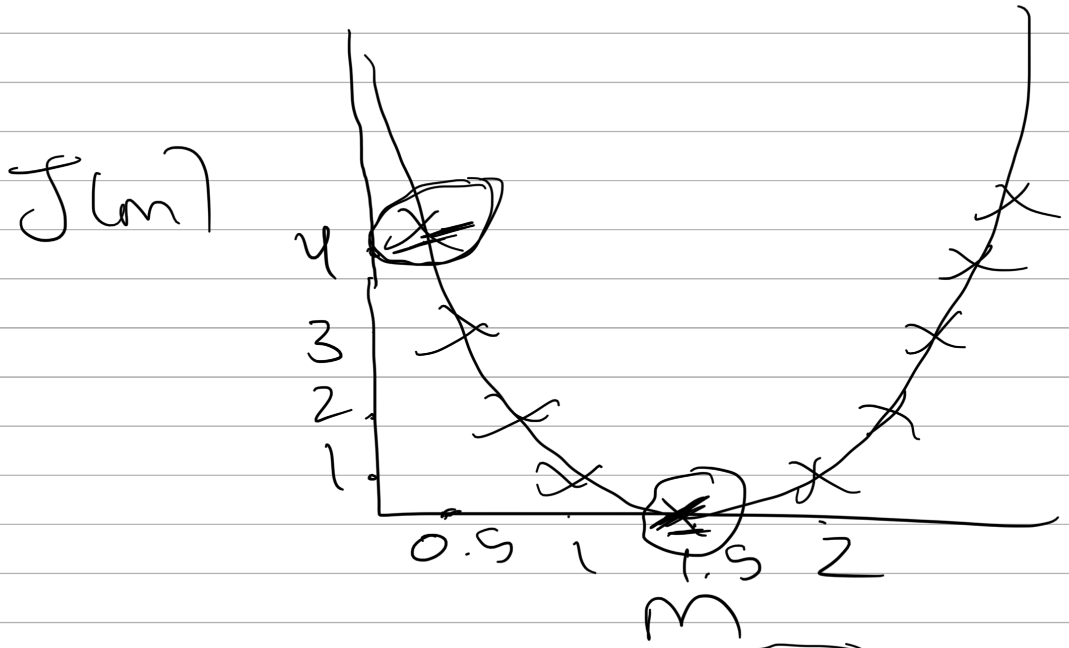
$x = 1$	y: 0.5
$x = 2$	1
$x = 3$	1.5

$$\frac{1}{2\times 3}\left((0.5-1)^2 + (1-2)^2 + (1.5-3)^2\right)$$

$$= 0.58$$

J

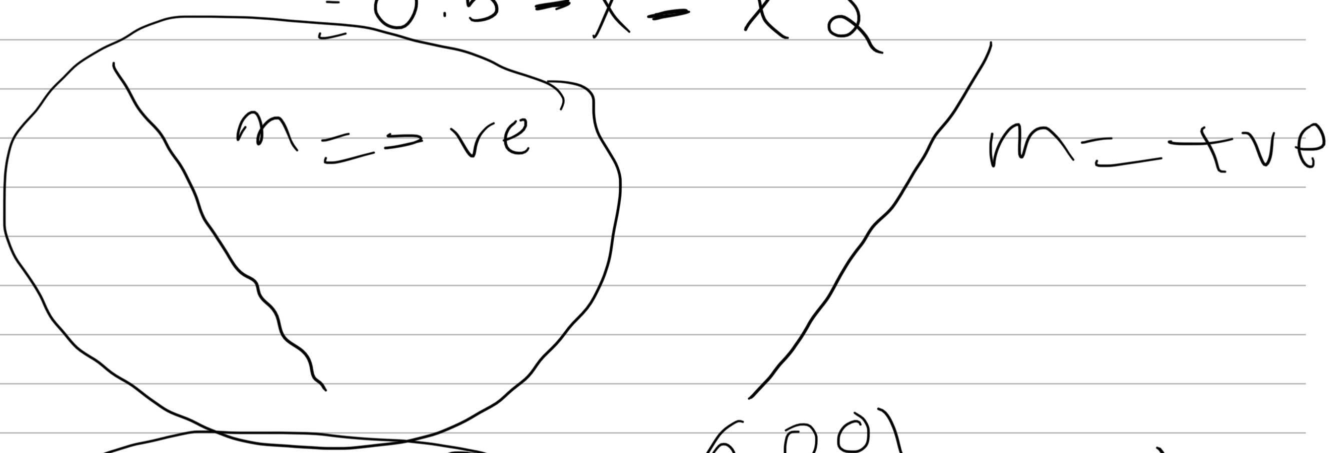
$m = 0.5$



Learning Rate

$$m_{new} = m - \left(\frac{\partial m}{\partial m}\right) \times \alpha$$

$= 0.5 - x - x\,\alpha$

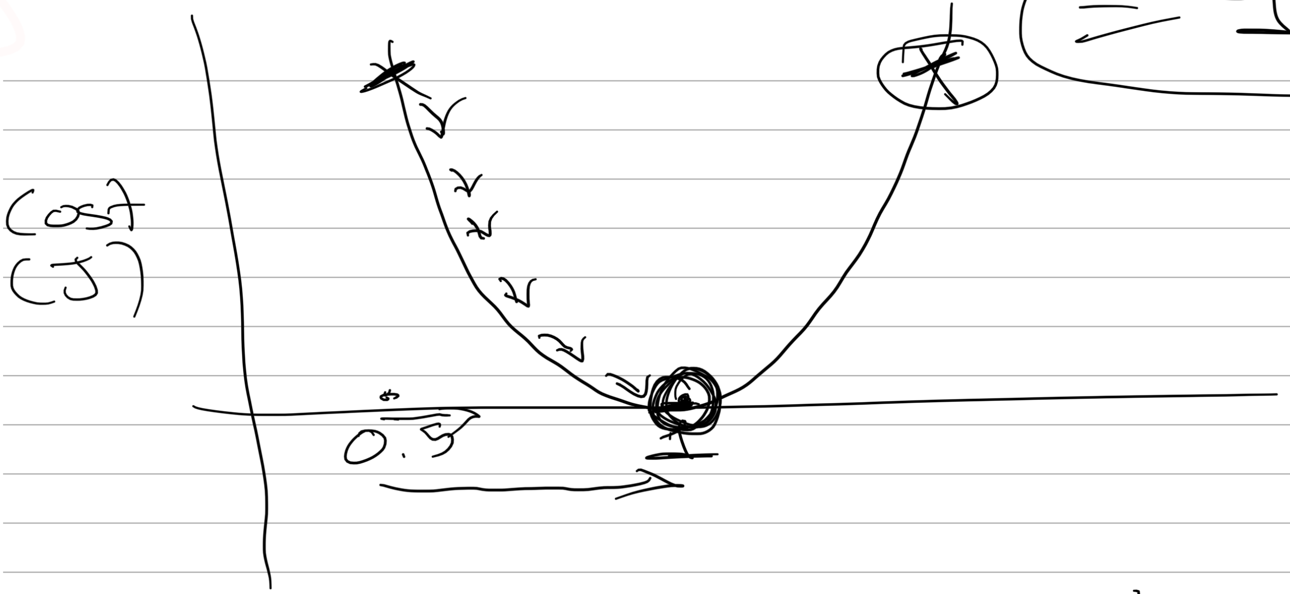


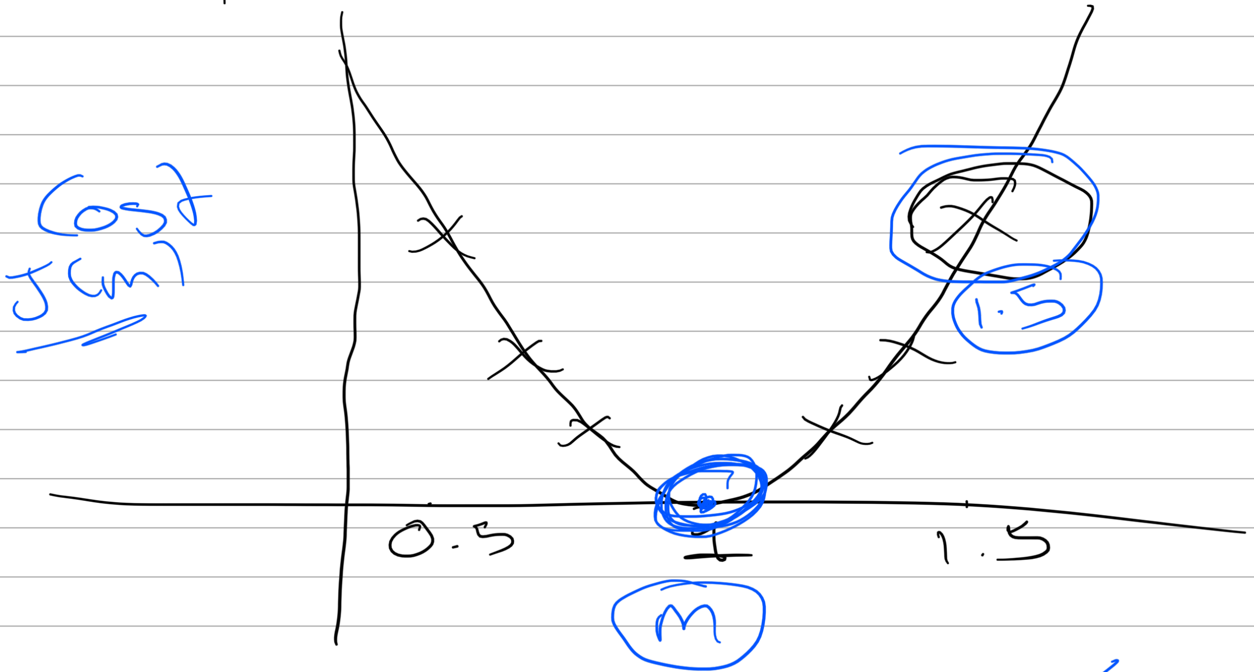
$m = 0.5 + \alpha$ — 0.001

$= 0.51$

$= 0.55$

$= 1$





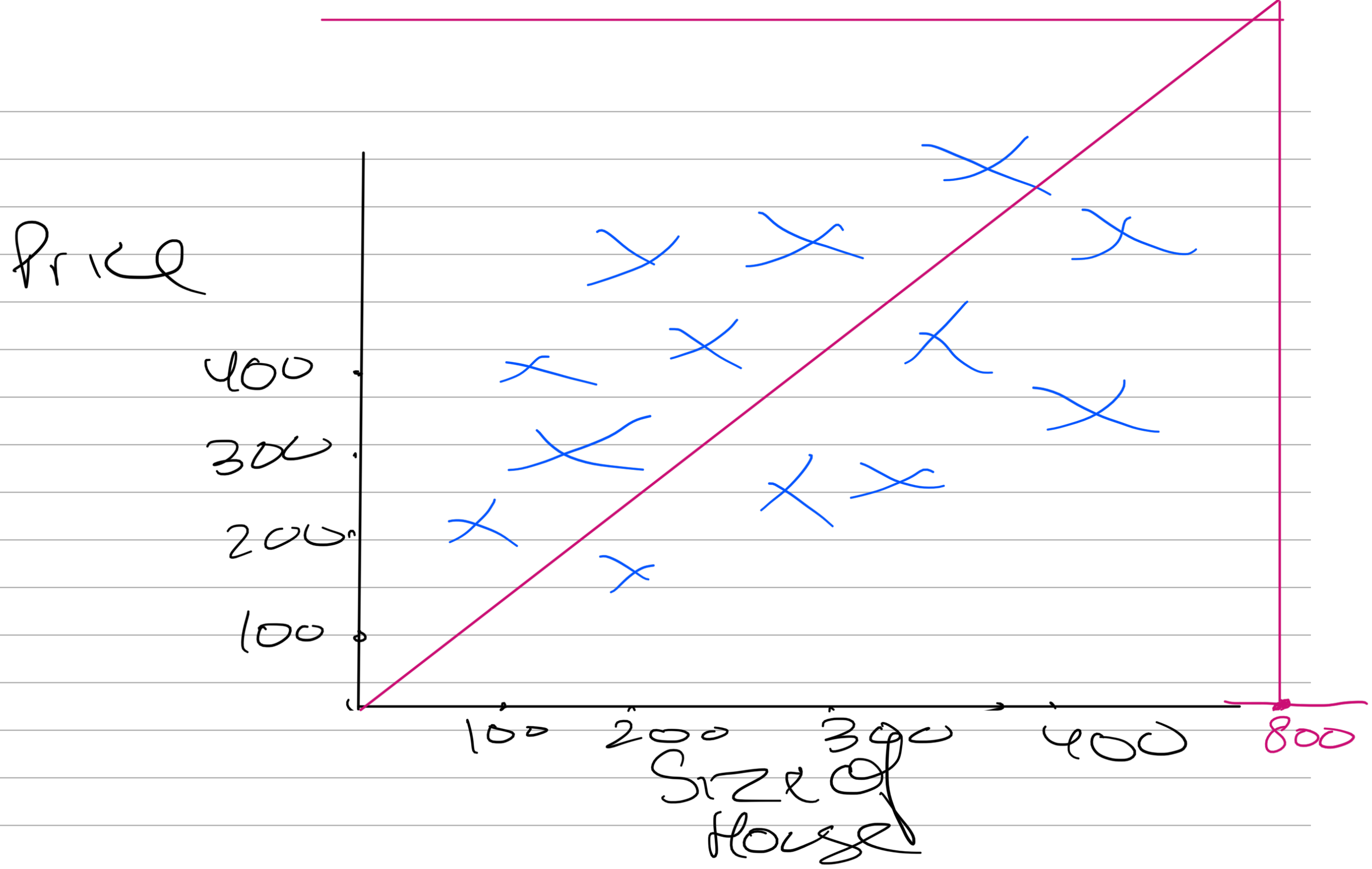
$m = +ve$

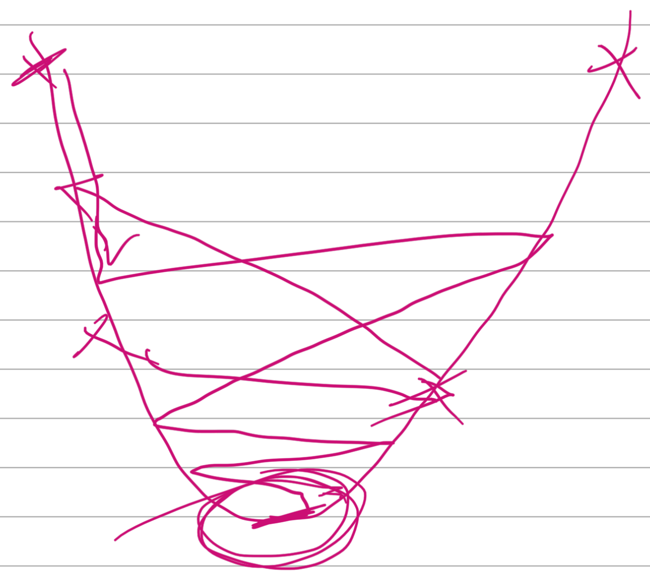
$$m_{new} = m_{orig} - \alpha \left(\frac{\partial m}{\partial m} \right)$$

$$= 1.5 - \alpha \times (+)$$

$$= 1.5 - \alpha$$

$\alpha = 0.01$


Price
400
300
200
100
100
200
300
400
800
Size of House



Sales

$y = mx + c$

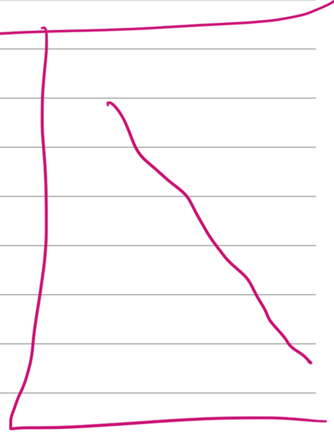
TV Radio

1. It states that the dependent variable Y should be linearly related to independent variables

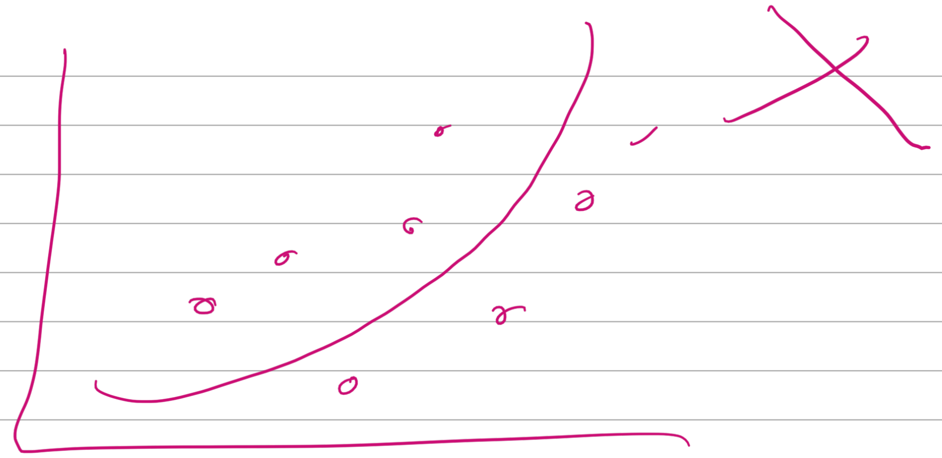
Y

x →

Y



x →



There should be no outlier

No high CoRelation